Niveau(x) d'enseignement: 0 Numéro EDUQ : 8770951
Centre d'intérêt: 540

Manuals on educational evaluation : manual two : glossary / Direction du développement de l'évaluation, Direction générale de l'évaluation et des ressources didactiques. -- [Québec] : Ministère de l'éducation, 1987, c1986.
41 p. ; 26 cm.

Code: 16-7265A ; 8687-524.

ISBN 2-550-13420-6

Vedette(s)-secondaire(s): Québec (Province) Ministère de l'éducation. Direction du développement de l'évaluation • Québec (Province) Direction générale de l'évaluation et des ressources didactiques

Descripteur(s): Lexique • Terminologie • Evaluation en éducation

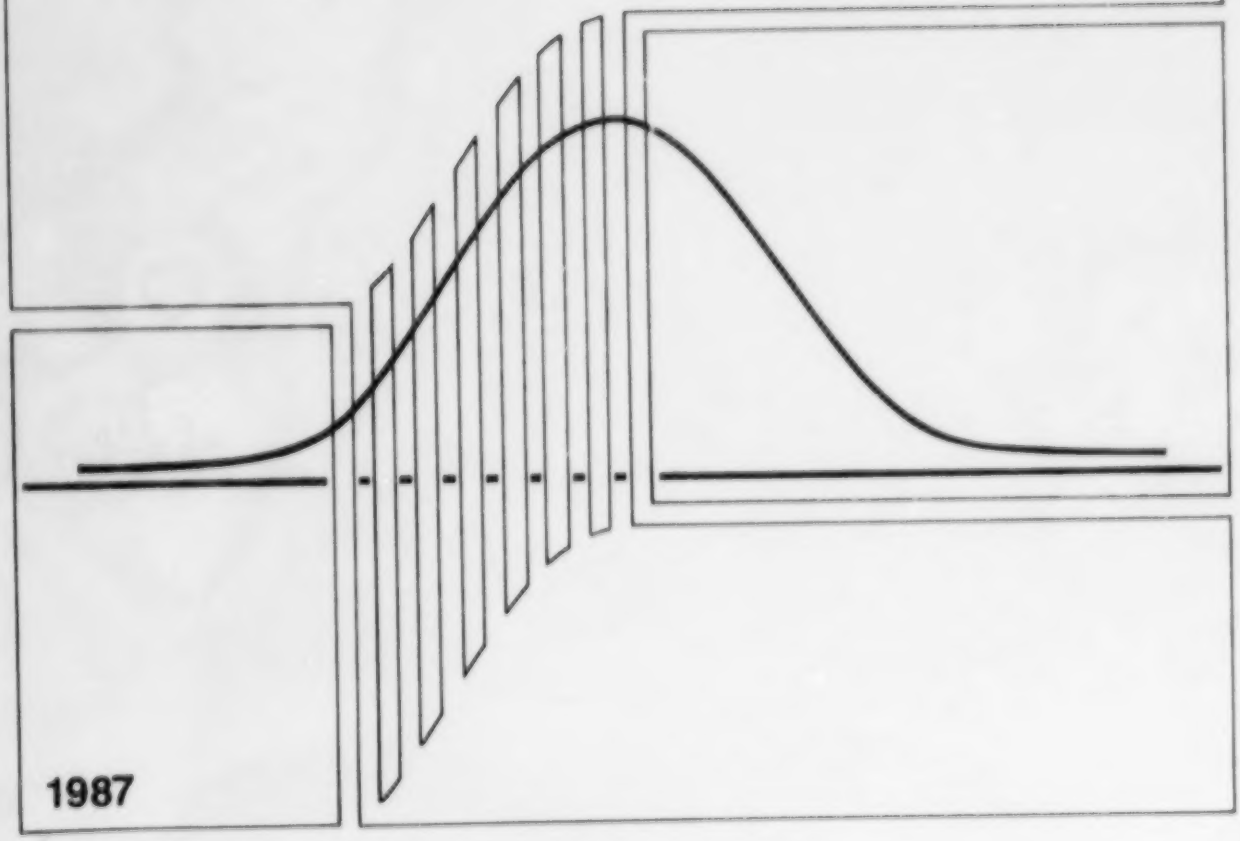
La terminologie de l'évaluation en éducation comprend surtout des termes ordinaires utilisés dans un sens spécifique. Ces termes ont été compilés et sont définis dans ce lexique. On y a ajouté quelques autres termes qui traitent plus spécifiquement d'évaluation ou de statistiques. Le but de ce lexique n'est pas de restreindre l'usage courant des termes en évaluation utilisés par les enseignants, mais bien de définir avec précision les concepts utilisés dans la série "Manuals on Education". Le lexique est structuré de façon à montrer les relations entre les concepts. Chaque définition est généralement suivie de la rubrique "Voir" sous laquelle on trouvera des mots utilisés dans la définition qui ont été définis ailleurs dans le lexique et, le cas échéant, des renvois. Deux index facilitent l'utilisation du lexique. L'index alphabétique donne la liste des concepts qui ont été définis et les synonymes les plus fréquemment utilisés. L'index thématique regroupe les termes qui se rapportent à une même question.

Type de document: d

EDUQ 8770951

Manuals on Educational Evaluation

Manual Two: Glossary



1987

Québec

EDUQ 8770951

Manuals on Educational Evaluation

Manual Two: Glossary

Direction du développement de l'évaluation
Direction générale de l'évaluation et des ressources didactiques

This publication is the work of the Direction générale de l'évaluation et des ressources didactiques du ministère de l'Éducation.

Edited by the Direction générale des publications gouvernementales du ministère des Communications.

© Gouvernement du Québec
Ministère de l'Éducation, 1986

ISBN 2 - 550 - 13420-6

Dépôt légal - quatrième trimestre 1986
Bibliothèque nationale du Québec

CONTENTS

PAGE

Foreword 1

Definitions 3

Alphabetical Index 35

Thematic Index 37

Acknowledgements 41

The terminology of educational evaluation consists, for the most part, of ordinary terms used in a specialized sense. This GLOSSARY is primarily concerned with the compilation and definition of such terms. Added to these are a few others specifically related to evaluation and statistics. The goal of this glossary is not to restrict current usage of some evaluation terms by educators in Quebec, but rather to define precisely the underlying concept used in the series "Manuals on Educational Evaluation".

The GLOSSARY is structured to show interrelationships among the different concepts. As a rule, each definition is followed by the heading "SEE" under which will be found: words used in the definition that receive their own definition in the GLOSSARY and, when appropriate, cross references. To facilitate the use of the GLOSSARY, two indices have been prepared. The Alphabetical Index includes the list of terms defined and the most frequently used synonyms. The "Thematic Index" groups terms related to a specific question.

Needless to say, each definition and cross reference involves choices which reflect a particular view of educational evaluation by specialists at the Ministère. This can be clearly seen, for example, by glancing at the synonyms for the expression "Standard of Performance."

As part of the Manuals on Educational Evaluation series, the GLOSSARY offers definitions derived from the review and analysis of educational evaluation documents published by the Ministère. When consulting it, the reader should not expect to find universal definitions, but rather a coherent set of terms and concepts related to a particular view of educational evaluation which is fostered by the Direction du développement en évaluation. As the Direction du développement en évaluation pursues its work of research and experimentation, this underlying view will no doubt evolve, and thus require revision of the definitions offered here.

A

ABILITY

SEE: • Skill

ACHIEVEMENT CRITERIA

SEE: • Standard of Performance

APPRAISAL

Procedure for estimating the quality of outcomes, based on specific criteria.

The same term is used to designate the results recorded on a rating scale.

SEE: • Criterion
• Learning
• Rating Scale

ASSESSMENT

Synonym: Evaluation

SEE: • Decision
• Educational Evaluation

ATTITUDE

The degree of an individual's like, dislike or ambivalence for an institution, situation, event or person.

N.B. This is a general definition. There are various notions of what an attitude is and each would require a separate definition.

ATTITUDE SCALE

A measure of the degree of an individual's like, dislike or ambivalence for an institution, situation, event, or person.

SEE: • Attitude

AVERAGE

The measure of central tendency of a group of scores. The sum of all raw scores on a given test divided by the number of scores.

The average is very much affected by the magnitude of extreme (very low or very high) scores.

SEE: • Central Tendency
• Mean
• Standard Deviation

B

BALANCED ASSESSMENT PROFILE

A collection of documents to be used by educators in order to diagnose the strengths and weaknesses of pupils having learning problems. These documents should be used in order to plan a specific remediation program for each student.

BEHAVIORAL OBJECTIVES

The specific objectives, expressed in terms of behavioral outcomes, to be achieved by a student within a specified program.

C

CEEB SCORES

A standard score scale developed by the College Entrance Examination Board and having a mean of 500 and a standard deviation of 100.

CENTILE

Synonym: Percentile

CENTRAL TENDENCY

The average or central score in a group of scores generally known as the mean or the median. The central tendency is a very important index to consider when interpreting results since this score represents the overall group performance.

SEE: • Mean
• Median

CONVERSION

Statistical method used by the MEQ in order to maintain the same examination difficulty level over the years. Conversion is meant to ensure that pupils will not be handicapped by an abnormally difficult examination. A more detailed definition is found in chapter four of the Administration Manual for the Certification of Secondary School Studies (16-7175A).

N.B. This term is not to be confused with "normalization."

SEE: • Moderation

CORRECTION KEY

Set of correct or admissible answers for the items of a test. List of elements expected in an answer.

This key, which is prepared during test construction, can also indicate the weight to be given to each element of a correct answer and to each test item, thus ensuring greater uniformity of marking.

Synonym: Correction Scale.

SEE: • Marking

CRITERION

A standard, by which performance is evaluated and marks are assigned.

SEE: • Appraisal
• Criterion-Referenced Evaluation
• Standard of Performance

CRITERION-REFERENCED EVALUATION

Evaluation using a test, examination or questionnaire that has been designed specifically for a criterion-referenced interpretation with very specific content specifications.

SEE: • Performance

CRITERION-REFERENCED INTERPRETATION

Interpretation of measurement results in terms of performance related to a task or set of tasks.

SEE: • Measurement
• Norm-Referenced Interpretation
• Performance

CUT-OFF SCORE

SEE: • Standard of Performance

DIRECTION DES COMMUNICATIONS

D

DECISION

Third step of the process of educational evaluation. It involves making a choice among various courses of action related either to learning progress or to recognition of learning achievements.

SEE: • Educational Evaluation

DEFINITION OF DOMAIN

A reference document which presents the essential elements of a program in Québec and is used primarily for the development of formative and summative evaluation instruments.

SEE:
- Behavioral Objectives
- Formative Evaluation
- Specification
- Summative Evaluation
- Table of Specifications

DERIVED SCORE

Raw score which has been modified by conversion, moderation or normalization.

Synonym: Modified Score.

SEE:
- Conversion
- Moderation
- Normalization
- Raw Score

DISPERSION

The extent to which a group of scores vary or deviate from the mean.

SEE:
- Average
- Mean
- Standard Deviation

DISTRACTOR

An incorrect but plausible response for a multiple-choice test item.

SEE:
- Multiple-Choice Item
- Plausibility

DOCIMOLOGIE (Derived from Greek dokimasia, a civil service examination in ancient Athens.)

Science of measurement and evaluation. It has as its major goal the development of tests, the interpretation of results and behavior of examinees, as well as the development and application of evaluation strategies.

N.B. This term is generally used in Europe and by Francophones in Québec.

E

EDUCATIONAL EVALUATION

A process which comprises:

a) gathering and interpreting (measurement) information;
b) making judgments;
c) deciding on a course of action (decision) based on the evaluation objectives set.

N.B. The series, Manuals on Educational Evaluation, is concerned only with the evaluation of students' school results and of some other facets of their general development whereas educational evaluation also includes evaluation of programs, personnel and institutions.

SEE: • Decision
• Judgment
• Measurement

EVALUATION TASK

That which is given to one or more students in order to measure their group performance or each student's performance and to judge it.

Synonym: Test situation.

SEE: • Appraisal
• Judgment
• Performance

EXAMINATION

A measurement instrument comprised of evaluation tasks given usually at the end of a school year or semester with uniform administration procedures in order to evaluate students' learning achievement in a given subject matter in order to grant credits.

N.B.: Unlike a test, an examination has not been field tested, nor validated.

SEE: • Evaluation Task
• Measurement Instrument
• Summative Evaluation

F

FORMATIVE EVALUATION

Process of ongoing evaluation which provides the teacher with information in order to help the student learn. This process is meant to ensure continuous learning progress through remedial or enrichment activities.

SEE: • Educational Evaluation
• Summative Evaluation

INDEX OF DIFFICULTY

A numerical value, usually in percentage indicating the level of difficulty of a test item. There are various ways of calculating such an index depending on the type of item analysis used and goal of the evaluation.

SEE:
- Item
- Item analysis

INDEX OF DISCRIMINATION

A numerical value indicating the extent to which a test item discriminates between students of higher or lower ability. This value is always found in an item analysis.

SEE:
- Item
- Item Analysis
- Measurement Instrument

INDEX OF INSTRUCTIONAL SENSITIVITY

A numerical value in decimals, with a range of .00 to 1.00 which indicates the performance changes (pretest - post-test) or differences between groups such as uninstructed - instructed. This term is also known as an index of discrimination (Berk, A Guide to Criterion - Referenced Test Construction, 1984).

Synonym: Index of Discrimination

INDICATOR

A reference point reflecting a performance or process which indicates that learning has taken place. Since learning is never measured directly, only inferred via indicators, the validity of these indicators is very important. The use of inference is thus required in defining what has been observed.

N.B. When used in the school report, the term refers to all signs indicating progress in student learning.

See: • Inference

INFERENCE

The cognitive process which consists of arriving at conclusions based on reason.

Since learning cannot be directly measured, inference is used to judge from observed performance that learning has taken place.

SEE: • Indicator

ITEM

A part of a measurement instrument generally known as a test or examination question. It may take the form of a problem to be solved, a question, or a task.

SEE: • Task

ITEM ANALYSIS

The process of determining the psychometric qualities of test items. This process consists of three phases:

a) review of the relevance between items and objectives, of the distribution of scores, and of some aspects of reliability and validity;

b) calculation of indices (average, standard deviation, index of difficulty, and of discrimination, measurement error, etc.);

c) decision respecting the revision, replacement or retention of items.

Item analysis is aimed at ensuring the highest possible psychometric quality for a measurement instrument.

SEE:
- Average
- Index of Difficulty
- Index of Discrimination
- Index of Instructional Sensitivity
- Item
- Measurement Instrument
- Plausibility
- Reliability
- Standard Deviation
- Validity

ITEM BANK

SEE:
- Item Pool

ITEM POOL

A collection of items systematically classified for easy access. For example, by index of difficulty, index of discrimination, by subject. The item pool can be used to prepare measurement instruments corresponding to specific needs.

J

JUDGMENT

Second step of the process of educational evaluation. It consists of giving an opinion on the learning of a student or of a class.

SEE: • Educational Evaluation

L

LEARNING

Knowledge, skills and attitudes that students can acquire or develop.

LEARNING PROCESS

"Process, producing more or less lasting effect, by which new behaviors are acquired or existing behaviors modified through interaction with the environment." (De Landsheere, Dictionnaire de l'évaluation et de la recherche en éducation, (1979).

N.B. This is a general definition; there exist various means of articulating the concept learning and each leads to a specific definition.

M

MARK

Alphabetical or numerical symbol representing the evaluation of a performance.

Synonyms: Grade, Raw Score

SEE:
- Appraisal
- Performance
- Raw Score

MARKING

Act of assigning a mark or score based on a correction key.

SEE:
- Correction Key
- Mark
- Raw Score

MEAN

The sum of all the raw scores on a given test divided by the number of scores.

SEE:
- Average
- Median

MEASUREMENT

First step in the process of educational evaluation. It consists of gathering, organizing, and interpreting pertinent data.

SEE:
- Criterion-Referenced Evaluation
- Educational Evaluation
- Norm-Referenced Evaluation

MEASUREMENT INSTRUMENT

A global term for any instrument used in education with a set of items systematically grouped to collect pertinent data for use in passing judgments and making decisions.

This term is used to refer to examinations, observation checklists, rating scales, tests, etc.

SEE:
- Examination
- Item
- Observation Checklist
- Rating Scale
- Test

MEDIAN

The score, in a distribution of scores, below and above which 50% of the scores fall.

MODERATION

"A mathematical operation by which the marks received at the Ministère from locally prepared examinations are modified in the light of the marks received on ministerial examinations. The goal of moderation is to reduce to a minimum the effects of certain variables on the marks resulting from locally prepared examinations."

Refer to the Administrative Manual for the Certification of Secondary School Studies (16-7175A) for a more detailed explanation of this subject.

N.B. Not to be confused with normalization.

SEE:
- Conversion
- Normalized Scores

MODIFIED SCORE

Synonym: Derived Score.

MULTIPLE-CHOICE ITEM

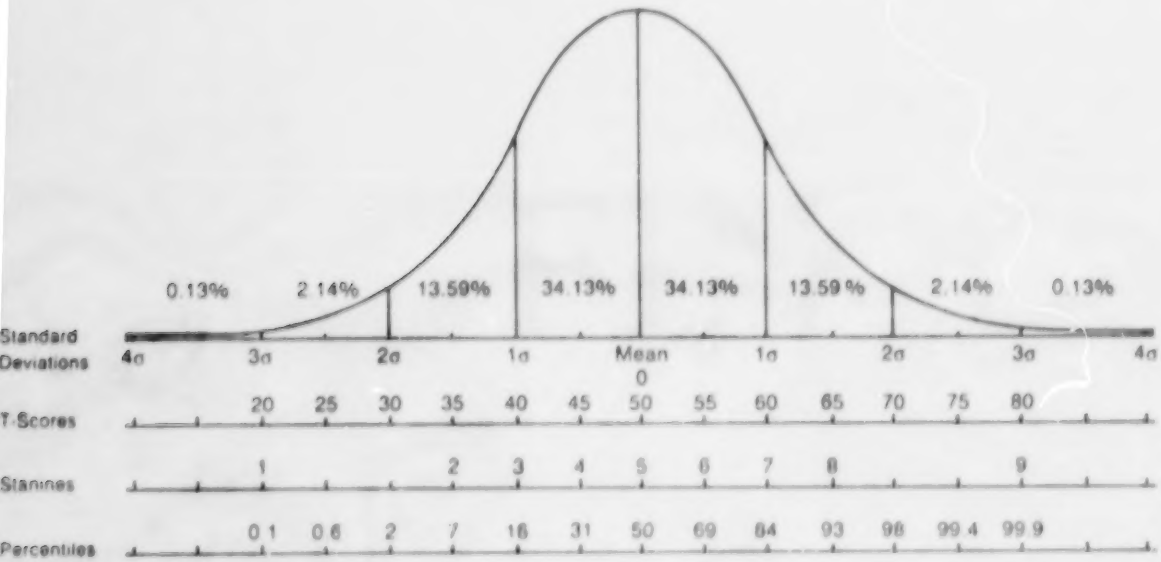
An item which requires the student to choose the correct or admissible response from a given set of responses. This type of item is usually found in a traditional objective test or examination.

N

NORMAL DISTRIBUTION

The frequency distribution of scores which is bell-shaped and symmetrical about the mean. The normal distribution is ASSUMED to apply to human abilities and achievements.

The illustration below indicates the essential features of this distribution.



NORMALIZATION

Transformation of scores which results in a frequency distribution identical to a so-called normal distribution.

After normalization, pupils will hold the same rank but their marks may be different.

SEE:
- Conversion
- Moderation
- Normal Distribution

NORMALIZED SCORES

Scores obtained by transforming raw scores in such a way that they are normally distributed about the mean. This transformation is done in order to be able to compare results from diverse groups.

NORM-REFERENCED EVALUATION

Assessing pupils by using a test, examination or questionnaire that has been designed specifically for a comparison between a student's raw score and some norm or reference group.

SEE:
- Criterion-Referenced Evaluation
- Norms

NORM-REFERENCED INTERPRETATION

Interpretation of a raw score in terms of a norm or reference group.

SEE:
- Criterion-Referenced Interpretation
- Measurement
- Norms

NORMS

Tables of scores and corresponding percentiles that describe the test performance of a reference group of students according to age, sex, school or any other relevant parameter.

Norms are the basis for the interpretation of raw test scores in a norm-referenced evaluation. Norms most currently used are the stanine and centile scales.

SEE:
- Centile Scale
- Stanine Scale

O

OBJECTIVE MEASUREMENT INSTRUMENT

A test, examination or questionnaire which is scored according to a predetermined correction key.

SEE: • Correction Key.

OBJECTIVITY

Characteristic of an evaluation process or of a measurement instrument resulting in a less biased interpretation of the test results.

An objective instrument will allow a pupil to obtain the same range of scores from different correctors or examiners.

OBSERVATION

Procedure by means of which the characteristics of an action or process are noted, while carefully avoiding any interference with the process or action assessed.

SEE: • Observation Checklist

OBSERVATION CHECKLIST

A device used to record the salient elements of an action, a product or a process; it furnishes a list of observable behaviors, as well as a means of recording observations.

Such checklists sometimes also include a marking system.

SEE: • Indicator
• Observation

P

PERCENTILE

The rank in percentage of a given raw score in relation to all other raw scores on the same test, examination or questionnaire. The xth percentile indicates the raw score below which x percent of examinee's scores fall.

PERFORMANCE

The demonstration, in a given situation, of behaviours involving knowledge, skills, or attitudes.

SEE: • Standard of Performance

PLAUSIBILITY

Essential characteristic of a distractor or wrong answer in a multiple-choice objective test.

SEE: • Distracter
• Item
• Multiple-Choice Test

PROMOTION

Act permitting a pupil who has met established requirements to continue studies in a higher class or cycle.

SEE: • Summative Evaluation

PSYCHOMETRIC

Term used in evaluation to denote the measurement qualities of an instrument. Some psychometric qualities of a test are, for example, difficulty level, reliability and validity levels, error of measurement, range, etc.

R

RANGE

The difference between the highest and lowest score on a given evaluation for a specific group.

RATING SCALE

A device for recording subjective estimates of scholastic achievements or of a performance. There are several types of rating scales; two examples are given below:

Verbal Scale

Unsatisfactory	Weak	Satisfactory	Good	Very Good

Descriptive Scale

Student answers without thinking, strays from the subject.	Student usually gives pertinent response, only occasionally off the subject.	Student almost always makes relevant comments.

SEE: • Appraisal

RAW SCORE

The total of a student's correct answers on a test, examination or questionnaire.

SEE: • Conversion
• Derived Score
• Modified Score

RECORDING

Act of writing down a student's achievement on an official document.

SEE: • Reporting

RELIABILITY

Characteristic of a measurement instrument which ensures that a student's rank order will remain stable over time.

The higher the reliability index of an instrument the less likelyhood of measurement error, thus the test scores represent more accurately a student's true ability level.

There are several types of reliabilities each with its shortcomings and strengths. The following are some important reliability coefficients.

- RELIABILITY: KUDER-RICHARDSON

A reliability coefficient which indicates the internal consistency of each test item within a given instrument. Kuder and Richardson, two psychometricians, devised a method to compute this relationship between each test item, as a result this coefficient is named after them. It should be noted that it is the most common published index of reliability in item analysis.

- RELIABILITY: PARALLEL FORMS

A reliability coefficient which indicates the degree of relationship between two evaluation instruments which have a common goal, the same content, and are of equal length.

- RELIABILITY: SPLIT-HALF

A reliability coefficient which indicates the internal consistency of an instrument by statistically dividing the test in half and computing the relationship between the two. A high coefficient (.80) indicates a high relationship, thus a high internal consistency.

- **RELIABILITY: TEST-RETEST**

  A reliability coefficient which indicates the degree of relationship between students' scores obtained in two sessions with at least three months interval between sessions. This coefficient reflects the stability of the instrument over time.

REPORTING

Process of providing information to educators, students and parents on a student's scholastic achievement. This process may take the form of an oral report, a printed report card including marks and or comments, or a computer print-out.

S

SCHOOL REPORT

"Document used to assign and report results of pupils' learning and school attendance. "(Regulation Respecting the Basis of Elementary School and Preschool Organization, Document #16-0063-01A p. II and Regulation Respecting the Basis of Secondary School Organization, Document #16-0062-01A.) p. II.

N.B. Research and pilot studies on the various methods of reporting school results are currently underway at the Ministère. This version of the GLOSSARY thus gives only a general idea of what the school report entails.

SEE: • Recording
• Reporting

SKILL

Set of behaviors related to the performance of a specific task in a given situation. The task in question may involve cognitive, affective or psychomotor abilities.

Synonym: Ability.

SEE: • Learning
• Performance

SPECIFICATION

A detailed description of the content of a test, examination or questionnaire. It is the blueprint of any evaluation instrument and is used to evaluate the instrument in question, and is also used as a guide in the development process of all evaluation instruments.

SEE: • Table of Specifications

STANDARD DEVIATION

A widely used measure of variability indicating the degree of dispersion of a set of scores around their average.

SEE: • Average
• Raw Score

STANDARDIZATION

Establishment of a set procedure for administering, scoring, and reporting results of a test, examination or questionnaire.

SEE: • Normalization

STANDARD OF PERFORMANCE

Level at which a specific performance is considered to be successful. This can be indicated by a score or a qualitative description obtained on the basis of an explicit criterion.

Synonyms: Criteria of Achievement
Cut-Off Score
Pass Score
Performance Criteria

SEE: • Criterion
• Criterion-Referenced Interpretation
• Performance

STANDARD SCORES

A group of scales such as CEEB scores and stanine scores having a desired mean and standard deviation.

STANINE

A standard score scale having a mean of 5, a standard deviation of 2 and a range from 1 to 9.

SEE:
- Mean
- Range
- Standard Deviation
- Standard Scores

SUBJECTIVE-RESPONSE ITEM

Item for which pupils must formulate their own response.

Types of items: Short-answer, essay, listing.

SEE:
- Item
- Multiple-Choice Item
- Task

SUMMATIVE EVALUATION

Evaluation at the end of an instructional unit, program, semester or year in order to determine the degree of learning by a student. This type of evaluation is primarily for promotion and certification by the Ministère and thus, assesses skills and knowledge judged essential for promotion and certification.

SEE:
- Educational Evaluation
- Formative Evaluation

T

TABLE OF SPECIFICATIONS

Organized presentation of the concepts or abilities which an evaluation instrument is supposed to sample. Included in this table is the relative importance of each concept and ability within the instrument.

SEE:
- Definition of Domain
- Specifications
- Taxonomy
- Weight

TASK

An activity that a student must do within the context of a learning or evaluation situation.

TAXONOMY

Hierarchical classification of cognitive, affective, or psychomotor behaviors, established on the basis of one or several criteria (complexity, internalization, etc.).

Several taxonomies have been developed, among which the most commonly used in Québec have been those of B.S. Bloom, R. Krathwohl and A. Harrow. However, the taxonomies of E. Simpson and L. D'Hainault are now enjoying increasing use.

SEE:
- Specifications
- Table of Specifications

TEST

Measurement instrument having proven psychometric qualities, and for which the degree of reliability and validity of results has been determined.

SEE:
- Measurement Instrument
- Reliability
- Standardization
- Validity

TEST BANK

A collection of tests systematically classified for easy retrieval according to their goals, psychometric qualities and subject matter.

SEE:
- Item Bank
- Test
- Test Battery

TEST BATTERY

A series of tests measuring different abilities or skills but standardized on the same population, thus allowing the comparison of test results for one person.

V

VALIDITY

Characteristic of a measurement instrument which indicates what this instrument is supposed to measure.

There are four major types of validities used to evaluate the relevance of an instrument for a specific goal. All the following validities are important to consider, however depending on the use of the instrument, the goal of the evaluation procedure and the type of decisions to be made, one type of validity may be more relevant than another. For example, in criterion-referenced summative evaluation the content validity of the instrument used is most important while predictive validity is of secondary importance.

SEE:
- Criterion-Referenced Evaluation
- Objectivity
- Reliability
- Summative Evaluation

- VALIDITY: CONCURRENT

Degree of correlation between results obtained by a group of students on two instruments A and B, where the validity of B is already established and accepted as a bench mark. The two instruments must measure the same or closely related subject matter.

- VALIDITY: CONSTRUCT

Degree to which scores from a test or examination permit solid inference about underlying traits. Construct validity is never determined by one study or analysis but rather by a continuous process of marshalling evidence on the relevance of the results.

- **VALIDITY: CONTENT**

Degree to which the content of the instrument represents what should be evaluated that is, specific learning based on a set program, and/or affective traits, and/or well defined psychomotor abilities. This validity is derived from a detailed analysis of the instrument's content by experts in the subject matter.

SEE: • Specification
• Table of Specifications

- **VALIDITY: PREDICTIVE**

Degree of correlation between results obtained on a test by a group of students and their subsequent performance. This validity is particularly relevant in aptitude testing, in selection and in employment.

W

WEIGHT

Importance given to each objective, item or score, indicating its relative importance within the total instrument.

SEE:
- Items
- Raw Scores
- Table of Specifications

ALPHABETICAL INDEX

A

Ability
Achievement Criteria
Appraisal
Assessment
Attitude
Attitude Scale
Average

B

Balanced Assessment Profile
Behaviorial Objectives

C

Centile
Central Tendency
Conversion
Correction Key
Criterion
Criterion-Referenced Evaluation
Criterion-Referenced Interpretation
Cut-Off Score

D

Decision
Definition of Domain
Derived Score
Dispersion
Distracter
Docimologie

E

Educational Evaluation
Evaluation Task
Examination

F

Formative Evaluation

I

Index of Difficulty
Index of Discrimination
Index of Instructional Sensitivity
Indicator
Inference
Item
Item Analysis
Item Bank
Item Pool

J

Judgment

L

Learning
Learning Process

M

Mark
Marking
Mean
Measurement
Measurement Instrument
Median
Moderation
Modified Score
Multiple-Choice Item

N

Normal Distribution
Normalization
Normalized Scores
Norm-Referenced Evaluation
Norm-Referenced Interpretation
Norms

O

Objective Measurement Instrument
Objectivity
Observation
Observation Checklist

P

Percentile
Performance
Plausibility
Promotion
Psychometric

R

Range
Rating Scale
Raw Score
Recording
Reliability
Reliability: Kuder-Richardson
Reliability: Parallel Forms
Reliability: Split-Half
Reliability: Test-Retest
Reporting

S

School Report
Skill
Specification
Standard Deviation
Standardization
Standard of Performance
Standard Scores
Stanine
Subjective-Response Item
Summative Evaluation

T

Table of Specifications
Task
Taxonomy
Test
Test Bank
Test Battery

V

Validity
Validity: Concurrent
Validity: Construct
Validity: Content
Validity: Predictive

W

Weight

THEMATIC INDEX

School Results

Grade
Mark
Marking
Recording
Reporting
School Report

Statistical Procedures

Conversion
Moderation
Normalization
Standardization

Statistical Indices

Centile
Percentile
Range
Stanine
Standard Deviation

TYPES OF INSTRUMENTS MEASUREMENTS

Attitude Scale
Examination
Measurement Instrument
Observation
Rating Scale

Planning

Specification
Table of Specifications
Taxonomy
Weight

Collection of Items or Instruments

Test Battery
Items

Analysis of an Instrument of Measurement

Correction Key
Index of Difficulty
Index of Instructional Sensitivity
Item Analysis
Objectivity
Plausibility
Psychometric
Reliability
Validity: Concurrent
Validity: Construct
Validity: Content
Validity: Predictive

TYPES OF EDUCATIONAL EVALUATION

Formative Evaluation
Summative Evaluation

INTERPRETATION OF DATA

Achievement Criteria (SEE: Standard of Performance)
Criterion
Criterion-Referenced Interpretation
Cut-Off Score (SEE: Standard of Performance)
Evaluative Task
Indicater
Norm-Referenced Interpretation
Performance Criteria (SEE: Standard of Performance)
Specification
Standard of Performance

PEDAGOGICAL EVALUATION

Appraisal
Assessment
Decision
Judgment
Measurement
Observation

TYPES OF DECISION

Decision
Promotion

The GLOSSARY is one of five manuals published by the Direction du développement en évaluation. It pertains to the specific definitions of current evaluation and statistical terms used by Québec educators. Since these manuals are produced for educators working in English schools, we have taken the liberty to adapt the translated documents in order to render them more relevant.

The remaining four manuals in the series Manuals for Educational Evaluation will be translated during the 1986 and 1987 school year and sent to all English schoolboards. This series of reference manuals is offered to all Québec educators wishing to obtain current practical information with regard to educational evaluation. The contents of each manual are presented below.

Manual One - Basic Concepts in Educational Evaluation

This manual presents the elementary concepts that an educator should know and understand in order to undertake evaluation in schools. The three major stages in the evaluation process are described in detail with specific illustrations to aid teachers understand the importance of each stage.

The two chief evaluation approaches, formative and summative evaluation are presented with practical examples of the salient elements of each approach. This manual is an essential document for an educator wishing to undertake student evaluation.

Manual Two - Glossary

This manual presents the key terms and concepts that an educator should know and understand in order to undertake educational evaluation. The Glossary compiles in alphabetical order, with cross-references most of the current terms in evaluation and statistics. This manual is an easy to use document which will prove to be handy in the classroom, teachers' lounge and at the schoolboard offices.

Manual Three - Formative Evaluation

This manual will be of particular interest to teachers since the authors have taken great care in presenting all the relevant information with regard to formative evaluation. Teachers will find numerous practical examples of recommended formative evaluation techniques which can be used in a variety of classroom situations. This manual was expressly written with the teacher's needs in mind.

Manual Four - Summative Evaluation

The fourth manual in this series focuses on "end-of-year" evaluation. Summative and formative evaluation concepts, techniques and instruments are presented and compared. There are many useful examples of various technique used in summative evaluation which will help school administration plan their evaluation program. Although this manual deals with the "administrative" aspects of evaluation, teachers are urged to consult this manual in order to understand the differences and similarities between these two important educational evaluation approaches.

Manual Five - Educational Measurement Instruments

This manual deals with the selection of appropriate instruments for evaluation in education. The most used instruments in Québec are listed with their strengths and weaknesses. This manual provides information for the selection, purchase and development of instruments. As in all the manuals, there are many practical examples which a teacher can use in preparing a better evaluation instrument. The last section of this manual deals with practical advice to educators on how to compile and interpret data resulting from evaluating students in a classroom.

ACKNOWLEDGMENT

La Direction du développement en évaluation would like to acknowledge the invaluable cooperation of the "Groupe de liaison des docimologues en milieu scolaire" (G.L.D.M.S.) in the preparation of these manuals. We would like to thank the following persons for their active collaboration in the planning, writing and review of the French manuals.

Pierre Bérubé	C.S.R. de Chambly
Michel Labelle	C.S. Huntingdon
Pierre Lafleur	C.S.R. Blainville-Deux-Montagnes
Rita-Éva Lapointe	C.S. du Sault-St-Louis
Lise-Raymond Bérubé	Consultante
François Sarrazin	C.S. Baldwin-Cartier

We would particularly thank the following anglophone educators for their excellent work in translating and adapting the French documents.

Raymond Ballie	C.S. Laurenval
Christopher Dulude	C.S. Chomedey de Laval
Deborah Morrison	Concordia University
Marjorie Gawley	Ministère de l'Éducation

Pavel de Liamchin
Direction du développement en évaluation

Gouvernement du Québec
Ministère
de l'Éducation

16-7265 A
8687-524